

Method and apparatus for interference suppression in speech signals.

Veröffentlichungsnr. (Sek.) ☐ EP0411360, B1
Veröffentlichungsdatum : 1991-02-06
Erfinder : KAESSER JUERGEN DR (DE)
Anmelder : BLAUPUNKT WERKE GMBH (DE)
Veröffentlichungsnummer : ☐ DE3925589
Aktenzeichen:
(EPIDOS-INPADOC-normiert) EP19900113320 19900712
Prioritätsaktenzeichen:
(EPIDOS-INPADOC-normiert) DE19893925589 19890802
Klassifikationssymbol (IPC) : G10L3/02
Klassifikationssymbol (EC) : G10L21/02A1
Korrespondierende Patentschriften ☐ JP3070324
Zitierte Dokumente: US4723294

Bibliographische Daten

In a method and apparatus for interference suppression in speech signals for speech recognition purposes, the speech being converted along with other sound phenomena into an electric signal (summation signal) and the other sound phenomena consisting of noises and a sound phenomenon produced by electro-acoustic conversion of an audio signal, the noises are furthermore recorded and converted into a noise signal. The noise signal and the audio signal are passed via adaptive filters and subtracted from the summation signal. The resulting signal is compared with the noise signal on the one hand and with the audio

signal on the other hand. The results of the comparison are used to control the adaptive filters.



①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ **Offenlegungsschrift**
⑪ **DE 39 25 589 A 1**

⑳ Aktenzeichen: P 39 25 589.1
㉑ Anmeldetag: 2. 8. 89
㉒ Offenlegungstag: 14. 2. 91

㉓ Int. Cl. 5:
H 04 B 15/02
H 03 H 21/00
H 03 H 17/06
G 10 L 5/06

DE 39 25 589 A 1

㉔ Anmelder:
Blaupunkt-Werke GmbH, 3200 Hildesheim, DE

㉕ Vertreter:
Eilers, N., Dipl.-Phys., Pat.-Anw., 3200 Hildesheim

㉖ Erfinder:
Kässer, Jürgen, Dr., 3201 Diekholzen, DE

Prüfungsantrag gem. § 44 PatG ist gestellt

㉗ Verfahren und Anordnung zur Störfreiung von Sprachsignalen

Bei einem Verfahren und einer Anordnung zur Störfreiung von Sprachsignalen zu Zwecken der Spracherkennung, wobei die Sprache zusammen mit anderen Schallereignissen in ein elektrisches Signal (Summensignal) umgewandelt wird und wobei die anderen Schallereignisse von einem durch elektro-akustische Wandlung eines Audiosignals entstandenen Schallereignis und Geräuschen bestehen, werden ferner die Geräusche aufgenommen und in ein Geräuschsignal umgewandelt. Das Geräuschsignal und das Audiosignal werden über adaptive Filter geleitet und von dem Summensignal subtrahiert. Das resultierende Signal wird einerseits mit dem Geräuschsignal und andererseits mit dem Audiosignal verglichen. Die Vergleichsergebnisse werden zur Steuerung der adaptiven Filter verwendet.

DE 39 25 589 A 1

Beschreibung

Die Erfindung geht aus von einem Verfahren nach der Gattung des Hauptanspruchs und umfaßt ferner eine Anordnung zur Durchführung des erfindungsgemäßen Verfahrens.

5 Zur Erkennung von Sprache sind verschiedene Verfahren bekannt, bei denen jedoch die Erkennungssicherheit durch Hintergrundgeräusche beeinträchtigt wird. Derartige störende Schallereignisse treten beispielsweise in einem fahrenden Kraftfahrzeug auf, in welchem außerdem noch das Autoradio eingeschaltet ist.

Aufgabe der vorliegenden Erfindung ist es, ein Verfahren zur Störbefreiung von Sprachsignalen und eine Schaltungsanordnung zur Durchführung dieses Verfahrens anzugeben, wonach die Sprachsignale einer geeigneten Spracherkennungsanordnung zugeführt werden können.

10 Das erfindungsgemäße Verfahren mit den kennzeichnenden Merkmalen des Hauptanspruchs hat den Vorteil, daß die durch auftretende störende Schallereignisse erzeugten Störsignale mit hoher Sicherheit unterdrückt werden. Die zur Durchführung des erfindungsgemäßen Verfahrens erforderlichen Anordnungen können mit Mitteln der modernen Halbleitertechnik in einfacher Weise realisiert werden.

15 Durch die in den Unteransprüchen aufgeführten Maßnahmen sind vorteilhafte Weiterbildungen und Verbesserungen der im Hauptanspruch angegebenen Erfindung und vorteilhafte Anordnungen zur Durchführung des erfindungsgemäßen Verfahrens möglich.

Ein Ausführungsbeispiel der Erfindung ist in der Zeichnung anhand mehrerer Figuren dargestellt und in der nachfolgenden Beschreibung näher erläutert. Es zeigt:

20 Fig. 1 eine schematische Darstellung einer erfindungsgemäßen Anordnung in einem Kraftfahrzeug und

Fig. 2 ein Blockschaltbild einer erfindungsgemäßen Anordnung.

Gleiche Teile sind in den Figuren mit gleichen Bezugszeichen versehen.

In Fig. 1 sind von einem Kraftfahrzeug lediglich je ein Teil des Fahrgastraums 1 und des Motorraums 2 dargestellt. Ein Mikrofon 3 ist im Fahrgastraum, beispielsweise oberhalb der Windschutzscheibe, angeordnet und nimmt Sprache, Fahrgeräusche und Schall von einem Autoradio 5 auf, der von einem Lautsprecher 4 abgestrahlt wird. Das vom Mikrofon 3 abgegebene Summensignal $\Sigma(t)$ wird einer Anordnung 6 zur Durchführung des erfindungsgemäßen Verfahrens zugeführt, an deren Ausgang 7 das weitgehend von Störungen befreite Sprachsignal $D_1(t)$ abnehmbar ist.

Ein weiteres Mikrofon 8 ist im Motorraum 2 angeordnet, nimmt Fahrgeräusche auf und leitet diese als Signal $F(t)$ an die Anordnung 6 weiter. Schließlich wird das dem Lautsprecher 4 zugeführte Audiosignal als Signal $A(t)$ ebenfalls der Anordnung 6 zugeleitet. Falls keine monaurale Wiedergabe beim Autoradio vorgesehen ist, können mehrere Audiosignale $A(t)$ abgenommen und der Anordnung 6 zugeführt werden.

Bei den folgenden Betrachtungen werden die jeweiligen akustischen und elektrischen Signale gleichgesetzt. Das infolge der Schallabstrahlung des Lautsprechers 4 vom Mikrofon 3 aufgenommene Signal unterscheidet sich vom Signal $A(t)$ durch eine zunächst unbekannte Übertragungsfunktion. Das auf das Mikrofon 3 einwirkende vom Lautsprecher 4 stammende Signal wird daher als $A_1(t)$ bezeichnet. In ähnlicher Weise unterscheidet sich das vom Mikrofon 3 aufgenommene Fahrgeräusch $F_1(t)$ von dem durch das Mikrofon 8 aufgenommenen Fahrgeräusch $F(t)$. Das Summensignal ergibt sich somit zu $\Sigma(t) = S(t) + F_1(t) + A_1(t)$.

Die Anteile des Summensignals $F_1(t)$ und $A_1(t)$ sind $A_1(t) = H_A \times A(t)$ und $F_1(t) = H_F \times F(t)$, wobei H_A und H_F die entsprechenden Übertragungsfunktionen sind und \times eine Faltung bedeutet.

Um die Anteile $F_1(t)$ und $A_1(t)$ aus dem Summensignal zu entfernen, ist eine Kenntnis der zunächst unbekannten Funktionen H_F und H_A erforderlich. Hierzu dienen adaptive Filter, die in der Anordnung nach Fig. 2 vorgesehen sind. Die Signale $\Sigma(t)$ und $F(t)$ werden von den Mikrofonen 3, 8 über geeignete Verstärker 11, 12 und Tiefpaßfilter 13, 14 Analog/Digital-Wandlern 15, 16 zugeführt. Das Audiosignal $A(t)$ wird ebenfalls über ein Tiefpaßfilter 17 einem Analog/Digital-Wandler 18 zugeleitet. Die Tiefpaßfilter begrenzen die Bandbreite der Signale auf einen Wert, der für das nachgeschaltete Spracherkennungssystem notwendig ist.

Ein erstes adaptives Filter 19 dient zusammen mit einem Korrelator 20 zur Ableitung des Signals $A_1(t)$ aus dem Signal $A(t)$. In einem Subtrahierer 21 wird dann das Signal $A_1(t)$ von dem Summensignal $\Sigma(t)$ subtrahiert, wodurch das Signal $D_1(t)$ entsteht.

50 Ein weiteres adaptives Filter 22 und ein Korrelator 23 sind zur Ableitung des Signals $F_1(t)$ aus dem Signal $F(t)$ vorgesehen. Mit Hilfe eines weiteren Subtrahierers 24 wird das Signal $F_1(t)$ von dem Signal $D_1(t)$ subtrahiert. Am Ausgang 7 steht somit das Signal $D_2(t)$ zur Weiterleitung an ein Spracherkennungssystem zur Verfügung.

Das adaptive Filter 19 ist ein nichtrekursives Filter der Ordnung ungefähr 100 mit der Zielfunktion

$$55 \quad Z_1(t) = (\Sigma(t) - H_{AZ} \times A) \blacksquare A(t) = 0 \text{ für } H_{AZ} = H_A, \quad (1)$$

wobei \blacksquare die Korrelationsfunktion bedeutet. Die Zielfunktion nutzt aus, daß für das richtige H_{AZ} das Signal $(S(t) + F_1(t))$ unkorreliert zum Audiosignal $A(t)$ ist.

Die Berechnung der Filterkoeffizienten im einzelnen kann nach Standardverfahren der digitalen Signalverarbeitung erfolgen, wie sie beispielsweise im Aufsatz "Adaptive Noise Cancelling: Principles and Applications", Proceedings of the IEEE, Vol. 63, No. 12, December 1975, Seiten 1692 bis 1716 beschrieben sind. Die adaptiven Filter und die Korrelatoren können mit geeignet programmierten Signalprozessoren realisiert werden.

Das Ausgangssignal $D_1(t)$ des Subtrahierers 21 wird sowohl im Korrelator 20 verwendet als auch dem weiteren Subtrahierer 24 zugeführt. Es ergibt sich zu $D_1(t) = \Sigma(t) - H_{AZ} \times A(t)$.

65 Zur Bestimmung von H_F wird das bereits von den Audiosignalen befreite Signal $D_1(t)$ verwendet, welches sich aus Fahrgeräusch und Sprache zusammensetzt. Diese sind ebenfalls unkorreliert, so daß sich das für das Audiosignal verwendete Verfahren auch für das Geräuschsignal anwenden läßt. Dabei ist die Zielfunktion

$$Z_2(t) = [(D_1(t) - H_{FZ} \times F(t)) \cdot F(t) = 0 \text{ für } H_{FZ} = H_F(t)] \quad (2)$$

Das Ausgangssignal des Subtrahieres 24 wird dann

$$D_2(t) = S(t) + (H_F - H_{FZ}) \times F(t) + [(H_A - H_{AZ}) \times A(t)] \times F(t) \quad (3)$$

Aus Gleichung (3) ist ersichtlich, daß sich das Ausgangssignal $D_2(t)$ umso mehr dem Signal $S(t)$ nähert, je besser die Annäherung von H_{FZ} an H_F bzw. H_{AZ} an H_A erfolgt. Auch bei einer nicht vollständigen Annäherung erfolgt eine deutliche Reduzierung des Störanteils.

Patentansprüche

1. Verfahren zur Störfreiung von Sprachsignalen zu Zwecken der Spracherkennung, wobei die Sprache zusammen mit anderen Schallereignissen in ein elektrisches Signal (Summensignal) umgewandelt wird und wobei die anderen Schallereignisse von einem durch elektro-akustische Wandlung eines Audiosignals entstandenen Schallereignis und Geräuschen bestehen, **dadurch gekennzeichnet**, daß ferner die Geräusche aufgenommen und in ein Geräuschsignal umgewandelt werden, daß das Geräuschsignal und das Audiosignal über adaptive Filter geleitet und von dem Summensignal subtrahiert werden, daß das resultierende Signal einerseits mit dem Geräuschsignal und andererseits mit dem Audiosignal verglichen wird und daß die Vergleichsergebnisse zur Steuerung der adaptiven Filter verwendet werden.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Geräusche an ihrem Ursprung aufgenommen werden.
3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß das Audiosignal, das Geräuschsignal und das Summensignal zuvor tiefpaßgefiltert und analog/digital-gewandelt werden.
4. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß in einem Kraftfahrzeug das Geräusch außerhalb des Fahrgastraums, die Sprache und die anderen Schallereignisse im Fahrgastraum aufgenommen werden und daß das Audiosignal einer Audiosignalquelle (Autoradio, Kassettenspieler, CD-Spieler) entnommen wird.
5. Anordnung zur Durchführung des Verfahrens nach Anspruch 1, dadurch gekennzeichnet, daß zum Vergleich des resultierenden Signals mit dem Audiosignal bzw. mit dem Geräuschsignal Korrelatoren (20, 23) vorgesehen sind.
6. Anordnung nach Anspruch 5, dadurch gekennzeichnet, daß das Summensignal von einem Mikrofon (3) im Fahrgastraum eines Kraftfahrzeugs über einen Verstärker (11) und einen Analog/Digital-Wandler (15) Mitteln (21, 24) zur Subtraktion anderer Signale vom Summensignal zuführbar ist, daß das Audiosignal über einen weiteren Analog/Digital-Wandler (18) einem ersten adaptiven Filter (19) und einem ersten Korrelator (20) zuführbar ist, welcher das erste adaptive Filter (19) steuert und welchem ferner die Differenz zwischen dem Summensignal und dem Ausgangssignal des ersten adaptiven Filters (19) zuführbar ist, und daß das Geräuschsignal von einem weiteren, außerhalb des Fahrgastraums angeordneten Mikrofon (8) über einen weiteren Verstärker (12) und einen dritten Analog/Digital-Wandler (16) einem zweiten adaptiven Filter (22) und einem zweiten Korrelator (23) zuführbar ist, welcher das zweite adaptive Filter (22) steuert und welchem ferner die Differenz zwischen dem Summensignal und mindestens dem Ausgangssignal des zweiten adaptiven Filters zuführbar ist.
7. Anordnung nach Anspruch 6, dadurch gekennzeichnet, daß den Analog/Digital-Wandlern (15, 16, 18) Tiefpaßfilter (13, 14, 17) vorgeschaltet sind.
8. Anordnung nach Anspruch 5, dadurch gekennzeichnet, daß die adaptiven Filter (19, 22) nichtrekursive Filter sind.
9. Anordnung nach Anspruch 8, dadurch gekennzeichnet, daß die Ordnung der adaptiven Filter etwa 100 beträgt.

Hierzu 1 Seite(n) Zeichnungen

